

# A Conceptual Development of Quench Prediction App build on LSTM and ELQA framework

Matej Mertik<sup>a</sup>, Maciej Wielgosz<sup>b</sup>, Andrzej Skoczeń<sup>c</sup>

<sup>a</sup>*The European Organization for Nuclear Research - CERN, CH-1211 Geneva 23 Switzerland*

<sup>b</sup>*Faculty of Computer Science, Electronics and Telecommunications,  
AGH University of Science and Technology, Kraków, Poland*

<sup>c</sup>*Faculty of Physics and Applied Computer Science, AGH University of Science and Technology,  
Kraków, Poland*

---

## Abstract

This article presents a development of web application for quench prediction in Technology Department, Machine Protection and Electrical Integrity group, Electrical Engineering (TE-MPE-EE) at CERN. The authors describe an Electrical Quality Assurance (ELQA) framework, a platform which was designed for rapid development of web integrated data analysis applications for different analysis needed during the hardware commissioning of the Large Hadron Collider (LHC). In second part the article describes a research carried out with the data collected from Quench Detection System by means of using an LSTM recurrent neural network. The article discusses and presents a conceptual work of implementing quench prediction application for TE-MPE-EE based on the ELQA and quench prediction algorithm.

**Keywords:** Software development, Data analysis, Web applications, Electrical Quality Assurance (ELQA), Large Hadron Collider, Deep Learning, Recurrent neural networks

---

## 1. INTRODUCTION

The Large Hadron Collider (LHC) is the world's largest and complex experimental facility ever created. It was built by the European Organization for Nuclear Research (CERN) between 1998 and 2008 in collaboration with over 10,000 scientists and engineers from over 100 countries, as well as hundreds of universities and laboratories. Its 27 kilometre circumference incorporates (among others) 1232 superconducting dipole magnets that lie beneath France – Switzerland border near Geneva.

Starting up this largest and complex experimental facility requires many teams of professionals and complex procedures due to its superconducting characteristic in the domain of cryogenics, electrical engineering, electronics, computer science, material science, physics and others. The TE-MPE-EE at CERN provides an important task by

---

*Email addresses:* [matej.mertik@cern.ch](mailto:matej.mertik@cern.ch) (Matej Mertik), [wielgosz@agh.edu.pl](mailto:wielgosz@agh.edu.pl) (Maciej Wielgosz), [skoczen@ftj.agh.edu.pl](mailto:skoczen@ftj.agh.edu.pl) (Andrzej Skoczeń)

hardware commissioning of the LHC machine, where all circuits are verified with many electrical measurements in the machine in different conditions during the complex procedure. The signals from measurements are stored in different systems at CERN, where data are available for their analysis. ELQA team started to develop a data analytic application to look through patterns implemented in the hardware commissioning in 2008 and 2014. As such a framework for prototyping data analysis apps was developed in the TE-MPE-EE.

The ELQA framework provides building blocks for designing web application for data analysis. In this paper we present a conceptual development of Quench Prediction Application from the research work on the prediction of quenches and its preliminary results. We show the concept of integration of

- (a) a research on a deep learning algorithm for quench prediction and
- (b) a research in software design for prototyping data analysis application for TE-MPE-EE.

The paper is structured as following. In first chapters there is a description of the ELQA framework and software libraries used by software design, then LSTM concept for predicting quenches with deep learning is described on the real data. In the last part we are presenting a way and a concept of implementing the Quench Prediction App for TE-MPE-EE at CERN.

## 2. ELQA FRAMEWORK

ELQA framework was developed for rapid design of data analysis application at TE-MPE-EE. Framework was designed for ELQA group where there is a need to address a vast amount of data that are collected in the hardware commissioning of the Large Hadron Collider (LHC). In the commissioning procedure the conditions of the circuits of the machine are verified and implemented by different measurements through all 8 sectors of the machine. Based on software design following requirements were identified carefully:

- (a) searching signal's similarities over the vast amount of tests,
- (b) visualizing trends of measurements through time domain in data,
- (c) developing corresponding data models for different flavors of ELQA measurements such as TP4 (Test Procedure 4), DOC (Dipole Orbit Correctors), MIC (Magnet Instrumentation Check),
- (d) developing different scenarios of visualization of different group of measurements.

Two additional software characteristics were specially addressed:

- (a) application should run in web environment and
- (b) framework should support high interaction for the user when exploring the data.

As such ELQA framework provide different components for building the web applications. There are three main libraries used based on which the concept was defined:

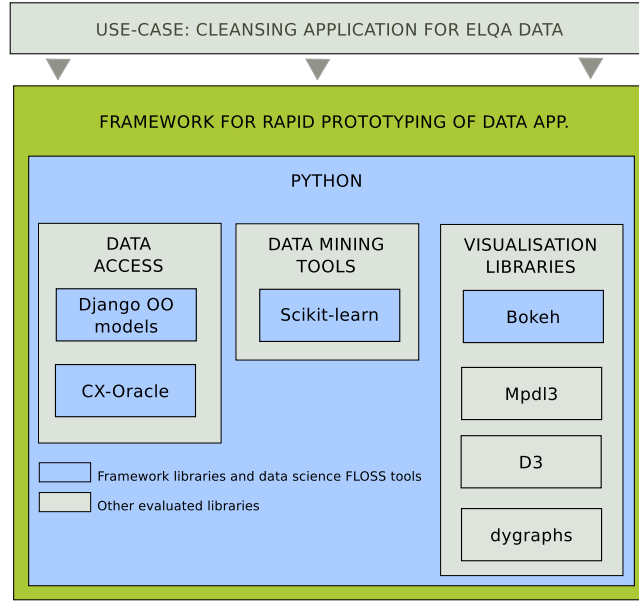


Figure 1: The components of the framework for rapid design of data applications

- (a) Bokeh Python interactive visualization library that targets modern web browsers for presentation [1], library was selected when testing the visualisation libraries and requirements [2],
- (b) Django high-level Python Web framework for Object Oriented paradigm to accessing relational dataset of ELQA [3], and
- (c) Scikit-learn kit, a Machine Learning library used for the implementation of the KDD algorithms [4].

Within these three main libraries other dependent libraries and components of CERN software infrastructure were integrated. The relationships between them are shown in Fig. 1.

Following subsections explain some detailed characteristic of the framework.

## 2.1. ELQA FRAMEWORK CHARACTERISTICS

### 2.1.1. Data models

Access to the data is addressed with the Object-Relational Mapping. Django framework handles this mapping providing all the functionality of a Structure Query Language (SQL). The architecture of Django is layered with

- (a) the bottom layer which is the ELQA database, followed by
- (b) access library that is responsible for communication between Python and the database through SQL statements, here we used a Python library cx.oracle [5] and

- (c) specific Django database back-end where any kind of database supported by Django (PostgreSQL, MySQL, SQLite and Oracle) is supported [3].

Above Django's layers a data model is then defined for accessing the data of ELQA. Data models can be efficiently programmed using the Django tools such as inspectdb for automatically defining the models from database tables, where relationships between the tables are separately defined.

#### 2.1.2. *Preprocessors*

Preprocessors are classes for analytical preprocessing of the signals within defined operators by the ELQA engineers. For the extraction of signals currently following basic statistics are used as extractors (and can be extended): average, minimal value, maximal value, skewness, kurtosis and properties of the linear regression applied to each signal such as slope and standard error. Scipy and numpy libraries [6, 7] are used by preprocessors implementation.

#### 2.1.3. *Data-miners*

Scikit-learn, a Machine Learning library written in Python is selected for data miners in the ELQA framework. It is built on NumPy, SciPy, and matplotlib and it is licenced under BSD open source licence [4]. Although there are many solutions and tools for data analysis available as open-source or commercial packages, some of the most used tools from the domains of Data Mining, Analytics, Big Data, and Data Science can be found on KDnuggets network maintained by a group of professionals and researchers [8], Scikit-learn was selected as most appropriate as is written in Python and importantly has a strong community around the library.

The data-miner layer defines general class for analyzer. Currently two different clustering algorithms are integrated from the scikit-learn such as K-means clustering [9] and DBScan clustering, a density-based spatial clustering of data with noise proposed by Martin Ester, Hans-Peter Kriegel, Jörg Sander and Xiaowei Xu in 1996 [10].

The analyzer class can be extended with other techniques from scikit-learn library or other libraries and algorithms in the framework. In the following section we present a concept of using LSTM-based solution for detecting and predicting quenches.

#### 2.1.4. *Visualizers*

Visualizers serve user interaction and visualization of the data. When searching for appropriate tools in communities many frameworks in different programming languages were reconsidered. Due to the required architecture (see other evaluated libraries on the Fig. 1) Bokeh was selected after studying the appropriate available solutions. Main reason was Bokeh's ability to act as a web server in order to expose a fully-fledged data visualization and analysis tool to multiple users [2].

### 3. CONCEPT OF USING THE LSTM-BASED SOLUTION

The idea behind using LSTM algorithm for quench detection and prediction is based on unique and very useful properties of this deep learning algorithm in time series modelling. It turns out that in some cases a currently used quench protection

solution [11, 12] can be supplemented with an additional monitoring layer which monitors a performance of the Quench Protection System.

One of the biggest challenges in implementing of the LSTM concept is the construction of the network architecture by choosing a right set of hyper parameters such as number of layers, neurons within each layer and input size. This is done in trial-and-error fashion which can be substantially leveraged by using an appropriate framework. Such a framework allows for visualization of the performance a given network instance as well as easy manipulation of its macro parameter values.

### 3.1. RECURRENT NEURAL NETWORKS

Phenomena occurring in a real world may be perceived in a spatial and temporal space. Since the neural models are supposed to mimic the real world the two different branches of neural networks evolved which address applications belonging to the aforementioned categories. Feed-forward and CNN networks are meant and most often used in applications of the spatial domain, whereas various kinds of recurrent models such as RNN, LSTM and GRU are especially useful in temporal modelling tasks.

Unlike traditional models that are mostly based on hand-crafted features deep learning neural networks can operate directly on raw data. This makes them especially useful in applications where extracting features is very hard and even sometimes impossible. It turns out that there are many fields of applications where no experts exist who can handle feature extraction or the area is simply uncharted and we do not know whether the data contains latent patterns worth exploring [13–15].

Foundations of the most currently used neural networks architectures were laid between 1960 and 1990. For almost the last two decades researchers were not able to take full advantage of these powerful models and there were claims that they are inherently ineffective. The whole machine learning landscape changed in early 2010 where deep learning algorithms started to achieve state-of-the-art results in a wide range of learning tasks. The breakthrough was brought about by several factors, among which computing power, deluge of widely available data and affordable storage are considered to be the critical ones. It is worth noting that in the presence of large amount of data the conventional simple linear models tend to under-fit or under-utilize computing resources.

Deep learning models can operate in both spatial and temporal domain. However, standard feed-forward and CNN neural networks are well suited for spatial domain [16–18], in contrast to recurrent neural networks which perform much better in the applications of temporal domain [19–21]. CNNs and feed-forward networks rely on the assumption of independence of data within training and testing set as presented in Fig. 2. This means that after each training item is presented to the model, the current state of the network is lost i.e. temporal information is not taken into account in training a model.

In a case of the independent data it is not an issue. But for data which contain crucial time or space relationships it may lead to a loss of majority of the information which is located in between steps. Additionally, feed-forward models expect fixed length of training vectors which is not always the case, especially when dealing with time domain data.

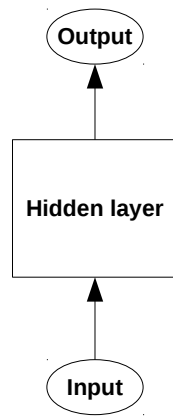


Figure 2: Architecture of standard feed-forward neural network

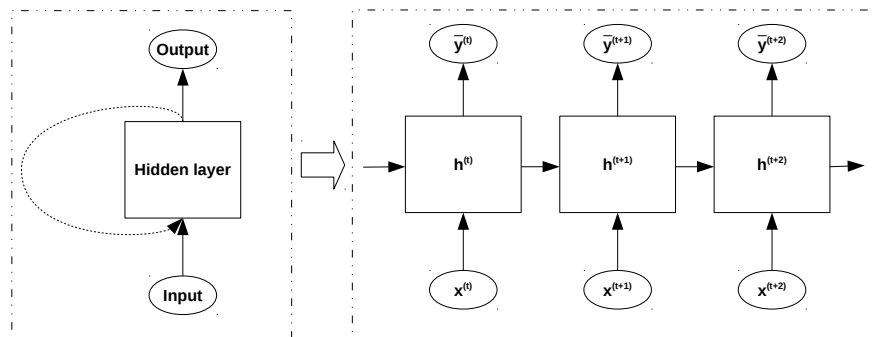


Figure 3: General overview of recurrent neural networks

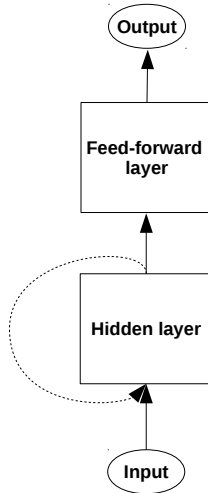


Figure 4: Mapping RNN outputs.

Recurrent neural networks (RNNs) are models with the ability to process sequential data as one element at a time. Thus they can simultaneously model sequential and time dependencies on multiple scales. Unfortunately, a range of practical applications of standard RNN architectures is quite limited. This is caused by the influence of a given input on a hidden and output layer during the training of the network. It either decays or blows up exponentially as it moves across recurrent connections. This effect is described as the *vanishing gradient problem* [22]. There were many unsuccessful attempts to address the problem before LSTM was eventually introduced by Hochreiter and Schmidhuber [23] which ultimately solved it.

Recurrent neural networks may be visualized both as a looped-back architectures of interconnected neurons which was presented in Fig. 2. Originally RNNs were meant to be used with single variable signals but they were also adapted for multiple stream inputs [22].

It is a common practice to use feed-forward recurrent layers together in order to map outputs from RNN or LSTM to the result space as presented in Fig. 4.

In practical applications, the LSTM has shown extraordinary ability to learn long-range dependencies as compared to standard RNNs. Therefore, most of the state-of-the-art applications use LSTM model [24].

The LSTM internal structure is based on a set of connected memory blocks as presented in Fig. 5. The memory blocks contain *memory cells* with loop-backed connections storing the temporal state of the network in addition to boundary units called *cell* which serve as an interface for information propagation within the network.

There are three different gates in each memory block:

- (a) *input* gate which controls input activations into the memory cell,
- (b) *output* gate controls cell outflow of activations into the rest of the network,





- (c) *forget* gate scales the internal state of the cell before summing it with the input through the self-recurrent connection of the cell. This enables gradual forgetting of a cell memory state.

Modern LSTM architectures may also contain *peephole connections* [25]. Since we do not use in our experiment it was not depicted in Fig. 5 and addressed in this description.

The output of each LSTM block is calculated according to the following set of equations:

$$g^{(t)} = \phi(W_{gx}x^{(t)} + W_{gh}h^{(t-1)} + b_g) \quad (1)$$

$$i^{(t)} = \sigma(W_{ix}x^{(t)} + W_{ih}h^{(t-1)} + b_i) \quad (2)$$

$$f^{(t)} = \sigma(W_{fx}x^{(t)} + W_{fh}h^{(t-1)} + b_f) \quad (3)$$

$$o^{(t)} = \sigma(W_{ox}x^{(t)} + W_{oh}h^{(t-1)} + b_o) \quad (4)$$

$$s^{(t)} = g^{(t)} \odot i^{(t)} + s^{(t-1)} \odot f^{(t)} \quad (5)$$

$$h^{(t)} = \phi(s^{(t)}) \odot o^{(t)} \quad (6)$$

While examining equations 1, 2, 3, 4 and 5 it may be noticed that instances for a current and previous time step are used for the value of hidden layer vector  $h$  as well as for internal state  $s$ . Consequently,  $h^t$  denotes a value of a hidden state vector at a current time step, whereas  $h^{t-1}$  refers to previous step. It is also worth noting the equations contain vector notation which means that they address whole set of LSTM cells. In order to address a single cell a subscript  $c$  is used as it is presented in Fig. 5 where for instance  $h_c^t$  refers to the value of hidden state within this particular cell.

LSTM based network learns when to let activation into the internal states of its cells and when to let the activation out. This is implemented with gating mechanism in which all the gates are considered as separate constituents of LSTM block with their own learning capability. This means that they adapt in a training process as separate units to preserve a proper information flow. When both the gates are closed, the activation is cut-off from the outside connections, neither growing or shrinking, nor affecting the output at intermediate time steps. In order to make it possible a hard sigmoid function  $\sigma$  was used which can output 0 and 1 as given by Eq. 7. This means that the gates can be fully open or closed.

$$\sigma(x) = \begin{cases} 0 & \text{if } x \leq t_l \\ ax + b & \text{if } x \in (t_l, t_h) \\ 1 & \text{if } x \geq t_h \end{cases} \quad (7)$$

In terms of the backward pass, so-called *constant error carousel* enables the gradient to propagate back through many time steps, neither exploding or vanishing [23, 24].

Table 1: Data sets used for training and testing the model

Data set	size [MB]
small	22
medium	111
large	5000

### 3.2. EXPERIMENTAL SETUP

Timber database stores a record of many years of the magnets activity. This is a huge amount of data with relatively few quench events. The first research question we had to address was related to a time span surrounding a quench event which should be taken into account for the LSTM model training. One day-long record of a single voltage time series for only one magnet occupies roughly 100 MB and there are several voltage time series associated with a single magnet [26] but ultimately we decided to use only one in our experiments, namely  $U_{res}$ .

Various kinds of magnets are used in LHC and they generate different number of quench events. It is highly beneficial from a perspective of our research to pick a magnet for which large possible number of quenches is provided. The longest history of quenches was provided for 600A magnets in Timber database. Therefore, we decided to focus our initial research on 600A magnets. Unfortunately, the data covering a quench periods of 600A magnets is very large i.e. an order of several gigabytes. However, as it was mentioned before, the activity record of superconducting magnets during operational time of LHC is composed mostly of sections of normal work and partially of quench events. Furthermore, Timber does not enable automated quench periods extraction, despite having many useful features for data preprocessing and information extraction.

Training deep learning models takes a long time even when fast GPUs are employed. Therefore, a choice of a size of a training data and a set of hyper-parameters of a model is critical in terms of the application performance. In course of experiments we discovered that it is faster to first train a model with a small data set and preliminarily adjust hyper-parameters and then tweak them on bigger data sets.

Consequently, we have created 3 different data sets: small, medium and the large ones as presented in Tab. 1.

The tests are in progress with a main goal of validating the LSTM-based setup capability of modeling superconducting magnets behavior. The resolution of the data acquired from Timber database is far too low to determine its ability to predict or detect quenches. This is will require using more data of higher resolution, such as post-mortem records, and is planned as the future work.

## 4. INTEGRATING LSTM MODEL INTO ELQA FRAMEWORK

Described LSTM concept developed at TE-MPE-EE can be efficiently used within ELQA framework for implementing quench prediction application for the end user.

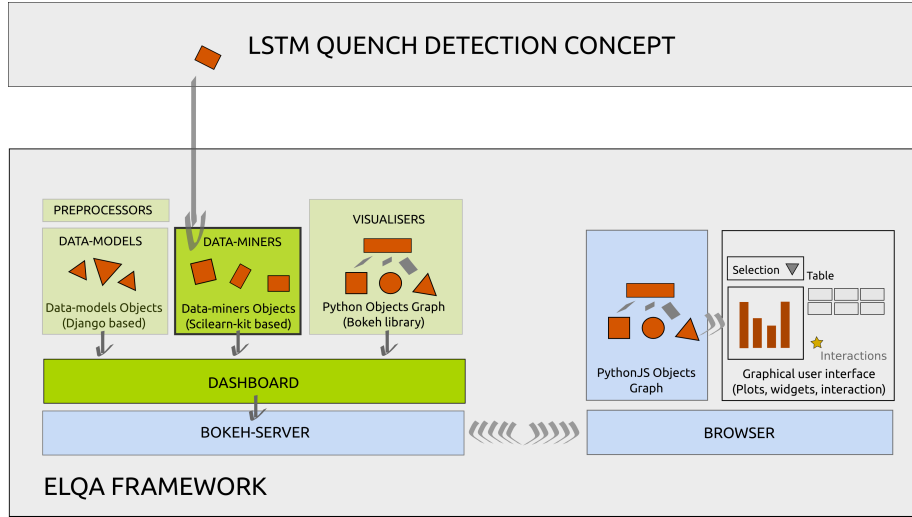


Figure 6: Integration of framework and LSTM

Following characteristics and key benefits are identified for joining these two pieces of research

- (a) efficient application of the research conducted on quench detection and prediction,
- (b) incorporating quench detection application into web environment
- (c) need of an efficient graphical user interface for end user,
- (d) availability of high user interaction experience on the developed application for quench detection.

Fig. 6 presents possible integration of the LSTM into the framework. It can be seen that LSTM concept should be wrapped by the new LSTM analyzer class and then used in the ELQA framework. The graphical user interface is a next step that should be designed through the definition of the Quench Prediction App Dashboard, a key part for any web apps developed for data analysis based on ELQA.

Following subsection provides an overview for integration process of the deep learning LSTM.

#### 4.1. INTEGRATING LSTM-BASED SOLUTION

Analyser class in ELQA is a general class where data miner algorithms are integrated from Scikit-learn data mining library. LSTM concept for quench prediction is implemented in Python at TE-MPE-EE and can be logical integrated into ELQA with an extended LSTM from the Analyser.

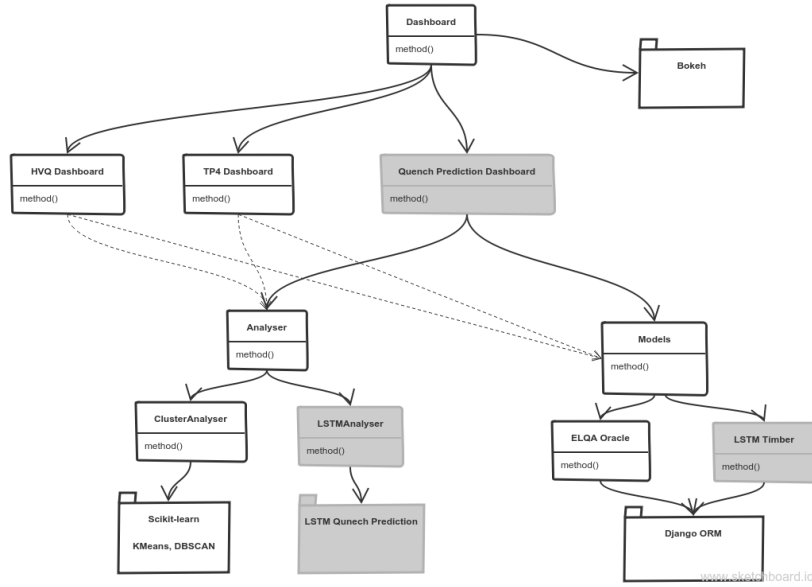


Figure 7: Integration of framework and LSTM, grey entities presents LSTM quench prediction parts in the ELQA framework

Accessing the data in Timber database will be extended by new Data model. ELQA framework allows to define new stack for integrating the access to the various datasets and stacks through Django library. Fig. 7 presents a possible integration.

When model and algorithm are incorporated into the ELQA classes, full development of the dashboard can be then applied through the framework for user interface and interaction developed for a Quench prediction app.

#### 4.2. IMPLEMENTING QUENCH PREDICTION DASHBOARD - APP DEVELOPMENT

With incorporated data model to access to Timber dataset and wrapped LSTM algorithm of Quench Prediction Software within Analyser class, efficient development of the web application for Quench prediction can be applied through programming of inherited LSTM dashboard. IN LSTM Dashboard widgets for Graphical User Interface are defined and then integrated through Bokeh library (input forms, plots), where data are applied to LSTM algorithm. Fig. 8 presents the concept of the Dashboard implementation by ELQA.

It can be seen (Fig. 8) that conceptual building blocks when programming Dashboard for quench prediction application follow objects that need to be defined in the Dashboard:

- (a) input widget for user interaction (lists, input boxes, sliders etc..),

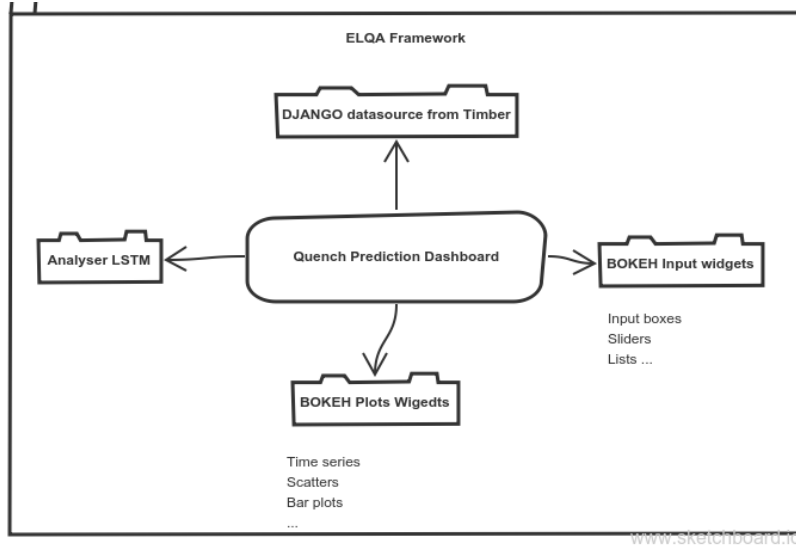


Figure 8: Dashboard design with ELQA for Quench Prediction app

- (b) plots widget for plotting series,
- (c) data sources from the data model (Django model for accessing Timber) and
- (d) LSTM algorithm as data-miner for quench prediction application.

## 5. CONCLUSIONS

We presented a concept paper of integrating the LSTM techniques for quench detection into the web application used the ELQA framework for prototyping data analysis application developed at TE-MPE-EE at CERN. Main characteristics of the framework were presented as the concept of LSTM for quench detection. We showed the mechanism and key benefits of joining these two entities for developing a quench prediction app for TE-MPE-EE at CERN. As explained, the LSTM concept can be integrated efficiently in order to provide an application that can be used by the engineers when providing hardware commissioning procedures of the LHC at CERN.

## Acknowledgments

The framework is being tested to prototype data application by the ELQA team. We would like to thank the entire ELQA team and TE-MPE-EE at CERN for their persistent invaluable help throughout the development process.

## References

- [1] Bokeh Development Team, Bokeh: Python library for interactive visualization (2014).  
URL <http://www.bokeh.pydata.org>
- [2] L. Barnard, M. Mertik, Usability of visualization libraries for web browsers for use in scientific analysis, *International Journal of Computer Applications* 121 (1).  
URL <http://research.ijcaonline.org/volume121/number1/pxc3904225.pdf>
- [3] Django Software Foundation. Django documentation [online] (2015).
- [4] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: Machine learning in python, *Journal of Machine Learning Research* 12 (Oct) (2011) 2825–2830.
- [5] A. Tuininga. cx\_Oracle [online] (2015).
- [6] E. Jones, T. Oliphant, P. Peterson, et al. SciPy: Open source scientific tools for Python [online] (2001–).
- [7] S. Van Der Walt, S. C. Colbert, G. Varoquaux, The NumPy array: a structure for efficient numerical computation, *Computing in Science & Engineering* 13 (2) (2011) 22–30. doi:10.1109/MCSE.2011.37.
- [8] G. Piatetsky-Shapiro. KDnuggets [online] (2015).
- [9] J. A. Hartigan, M. A. Wong, Algorithm AS 136: A K-Means Clustering Algorithm, *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28 (1) (1979) pp. 100–108. doi:10.2307/2346830.  
URL <http://www.jstor.org/stable/2346830>
- [10] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, in: *Kdd*, Vol. 96, AAAI Press, 1996, pp. 226–231.
- [11] L. Coull, D. Hagedorn, V. Remondino, F. Rodriguez-Mateos, LHC magnet quench protection system, *IEEE Transactions on Magnetics* 30 (4) (1994) 1742–1745. doi:10.1109/20.305593.
- [12] R. Filippini, B. Dehning, G. Guaglio, F. Rodriguez-Mateos, R. Schmidt, B. Todd, J. Uythoven, A. Vergara-Fernandez, M. Zerlauth, Reliability assessment of the LHC machine protection system, in: *Proceedings of the 2005 Particle Accelerator Conference, 2005*, pp. 1257–1259. doi:10.1109/PAC.2005.1590726.
- [13] C. Y. Chang, J. J. Li, Application of deep learning for recognizing infant cries, in: *2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, 2016, pp. 1–2. doi:10.1109/ICCE-TW.2016.7520947.

- [14] M. Cai, Y. Shi, J. Liu, Deep maxout neural networks for speech recognition, in: Automatic Speech Recognition and Understanding (ASRU), 2013 IEEE Workshop on, 2013, pp. 291–296. doi:10.1109/ASRU.2013.6707745.
- [15] L. Tth, Phone recognition with deep sparse rectifier neural networks, in: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 6985–6989. doi:10.1109/ICASSP.2013.6639016.
- [16] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (7553) (2015) 436–444, insight. doi:10.1038/nature14539.
- [17] C. Farabet, C. Couprie, L. Najman, Y. LeCun, Learning hierarchical features for scene labeling, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35 (8) (2013) 1915–1929. doi:10.1109/TPAMI.2012.231.
- [18] Y. LeCun, Deep learning convolutional networks, in: 2015 IEEE Hot Chips 27 Symposium (HCS), 2015, pp. 1–95. doi:10.1109/HOTCHIPS.2015.7477328.
- [19] J. Morton, T. A. Wheeler, M. J. Kochenderfer, Analysis of recurrent neural networks for probabilistic modeling of driver behavior, *IEEE Transactions on Intelligent Transportation Systems* PP (99) (2016) 1–10. doi:10.1109/TITS.2016.2603007.
- [20] F. Pouladi, H. Salehinejad, A. M. Gilani, Recurrent neural networks for sequential phenotype prediction in genomics, in: 2015 International Conference on Developments of E-Systems Engineering (DeSE), 2015, pp. 225–230. doi:10.1109/DeSE.2015.52.
- [21] X. Chen, X. Liu, Y. Wang, M. J. F. Gales, P. C. Woodland, Efficient training and evaluation of recurrent neural network language models for automatic speech recognition, *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 24 (11) (2016) 2146–2157. doi:10.1109/TASLP.2016.2598304.
- [22] A. Graves, *Supervised Sequence Labelling with Recurrent Neural Networks*, Springer Berlin Heidelberg, 2012. doi:10.1007/978-3-642-24797-2.
- [23] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.* 9 (8) (1997) 1735–1780. doi:10.1162/neco.1997.9.8.1735.
- [24] Z. C. Lipton, J. Berkowitz, C. Elkan, A critical review of recurrent neural networks for sequence learning (2015). arXiv:1506.00019.
- [25] K. Greff, R. K. Srivastava, J. Koutnk, B. R. Steunebrink, J. Schmidhuber, LSTM: A search space odyssey (2015). arXiv:1503.04069.
- [26] C. Roderick, L. Burdzanowski, G. Kruk, The CERN accelerator logging service - 10 years in operation: A look at the past, present, and future, *Proceedings of ICALEPCS2013* (2013) 612–614.